SHORT REPORT
Infectiousness of smallpox relative to disease age: estimates based on transmission network and incubation period

H. NISHIURA[1,2] AND M. EICHNER[1]*

[1] *Department of Medical Biometry, University of Tübingen, Tübingen, Germany*
[2] *Research Center for Tropical Infectious Diseases, Nagasaki University Institute of Tropical Medicine, Nagasaki, Japan*

SUMMARY

This study investigated the infectiousness of smallpox relative to the onset of fever using a likelihood-based estimation procedure based on the observed transmission network ($n = 223$) and on the distribution of the incubation period ($n = 379$). Who-infected-whom information enabled us to back-calculate the infectiousness by disease age, employing a step function model for infectiousness. Frequency of secondary transmissions was highest between 3 and 6 days after onset of fever, yielding an expected daily frequency of 20·6% (95% CI 15·1–26·4) of the total number of secondary transmissions, which is consistent with previous observations. The estimated cumulative frequency suggests that 91·1% of secondary transmissions occurred up to 9 days after onset of fever. The proposed method appeared to be useful for diseases with an acute course of illness, where transmission was not hampered by depletion of susceptible contacts.

When assessing the transmissibility of an infectious disease, it is important not only to measure the overall transmission potential [1], but also the illness-stage dependent infectiousness [2] to understand and clarify the effectiveness of control strategies. Transmission potential on a whole has been measured by the basic reproduction number, R_0, i.e. the average number of secondary cases arising from the introduction of an index case into a fully susceptible population [3, 4]. Since a reproduction number <1 indicates the eventual termination of an outbreak based on a threshold theorem, the knowledge of R_0 has been extremely useful in determining the required vaccination coverage and clarifying ecological and public health questions (e.g. periodicity of epidemics, average age at infection and heterogeneous transmissions) based on

population dynamics [5]. For simplicity, infectiousness is frequently assumed to be constant [6] and thus, the frequency of secondary transmissions by disease age (i.e. the time since onset of the disease) has rarely been evaluated. One of the rare exceptions is HIV infection whose infectiousness has been evaluated with regard to disease age, based on statistical analysis of partner studies [7]. The results suggest that the infectiousness of HIV is not constant, and this could largely influence the qualitative patterns of the epidemic [4, 8]. The knowledge of the disease-age-specific infectiousness is not only crucial for slowly progressing diseases but also for acute directly transmitted diseases, where targeted control measures, including isolation and contact tracing, have to be implemented as early as possible during the infectious period, and thus, their effectiveness largely depends on the time-course of infectiousness [9, 10].

One approach was to quantify how the pathogen load changes over time by using the most sensitive

* Author for correspondence: Dr M. Eichner, Department of Medical Biometry, University of Tübingen, Westbahnhofstr. 55, D-72070, Tübingen, Germany.
(Email: martin.eichner@uni-tuebingen.de)
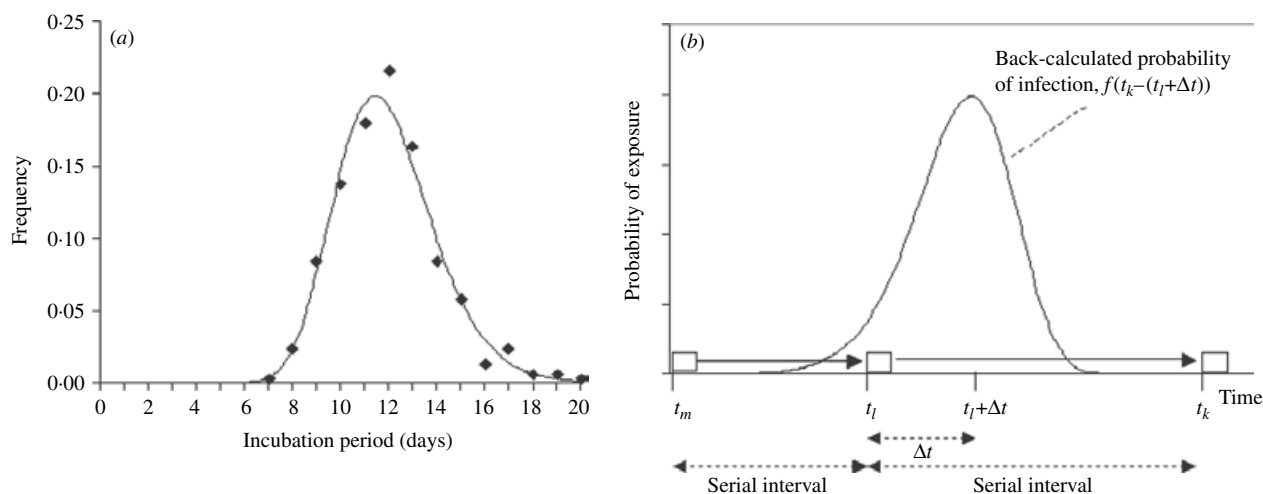
Fig. 1. Distribution of the incubation period and illustration of the estimation procedure. (*a*) Observed (◆) and fitted (—) lognormal distribution of the incubation period ($n = 379$; H. Nishiura & M. Eichner, unpublished observations). (*b*) Back-calculation of the transmission probability: case m infected case l who subsequently infected case k. Their times of onset are t_m, t_l and t_k, respectively. Using the difference of the disease onset (serial interval) $t_k - t_l$ together with the distribution of the incubation period, the disease-age-specific probability of transmission from case l to case k is obtained. □, Onset of cases; →, transmission.

microbiological techniques (e.g. real-time polymerase chain reaction) [11], but such observations are practically limited to the time after onset of symptoms and the pathogen load information can only be a useful measure of infectiousness if it is correlated with actual transmission. With regard to smallpox, several attempts have been made to measure the distribution of the virus-positive period [12, 13]. Despite rigorous studies, sample sizes were limited and, in particular, few samples could be obtained during the early stage of illness. Due to different laboratory techniques, the obtained viraemic periods differ widely. In this paper, we propose an epidemiological evaluation method, based on the distribution of the incubation period and the transmission network (who infected whom), to estimate how smallpox infectiousness varies over the course of illness.

The key information used in our approach includes the distribution of the incubation period and the transmission network. The incubation period, denoted by τ, is the time from infection to onset of disease (i.e. fever). The distribution has been fitted to a lognormal distribution, $f(\tau | \mu, \sigma)$, with the maximum-likelihood estimates of the mean and the standard deviation (S.D.) of $\log(\tau)$ being $\mu = 2\cdot47$ and $\sigma = 0\cdot17$, respectively (Fig. 1*a*; H. Nishiura & M. Eichner, unpublished observations). The transmission network is the observed chain of transmission that yields the information of

who infected whom. This type of information has already been explored to assess the number of secondary transmissions over the course of an epidemic [14] and to evaluate individual variations in transmission [15, 16], but it also enables us to obtain the serial interval, i.e. the time from symptom onset in a primary case to symptom onset in a secondary case [17, 18]. This study uses observed serial intervals in five smallpox outbreaks, in Glasgow (1950), Brighton (1950–1951), Industrial Pennines (1953), Bawku (1967) and Calcutta (1971–1972) [19–23]. A total of 223 serial intervals were identified. The mean (and median and S.D.) of these intervals was 16·0 (16·0 and 4·0) days. The minimum and maximum intervals were 6 and 37 days, respectively.

Figure 1*b* illustrates our estimation method using a chain of transmission of three consecutive cases, m, l and k. Based on observation, it is known that case m infected case l who subsequently infected case k. We denote the times of onset of these cases by t_m, t_l and t_k, respectively, where $t_k - t_l$ is the serial interval between cases l and k. Using this interval together with the distribution of the incubation period, the disease-age-specific infectiousness of case l is obtained. Assuming that the probability of observing transmission from primary case l to the secondary case k for each day can be extracted from the independently identically distributed incubation period, the probability density

that case k was infected Δt days after onset of case l is given by:

$$P_{l,k}(\Delta t)$$
$$= \begin{cases} \dfrac{f(t_k - (t_l + \Delta t))}{\int_{-x}^{t_k - t_l} f(t_k - (t_l + \tau))\mathrm{d}\tau}, & \text{if} -x \leqslant \Delta t \leqslant t_k - t_l, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Transmission from case l to k may have occurred before the onset of case l. Although one could consider any time of transmission dating back to t_m, we restrict the potentially contagious period to $x = 5$ days before onset of primary case l. As the incubation period lasts longer than 5 days in over 99·9% of cases, this assumption does not create major conflicts (e.g. we do not observe transmission from case l to k before m infected l). Some secondary cases were vaccinated before exposure or received post-exposure vaccination, but we assume that their vaccination did not significantly influence the length of the incubation period, because the cases did develop smallpox. Since the parameters of the incubation period distribution are known, eqn (1) can numerically be solved by discrete approximation of the incubation period distribution. Other underlying assumptions include the following: (i) the frequency of secondary transmission is independent of the overall transmission potential; although the contagiousness of vaccinated cases might have been lower than that of unvaccinated ones, we assume that their relative contagiousness in the different phases of disease was that of unvaccinated cases, (ii) human behaviour, spatial information and extrinsic factors (i.e. the implementation of control measures) can be ignored (for discussion on this point see below).

Using these assumptions, the disease-age-specific infectiousness can be estimated by using a likelihood-based procedure. We denote the number of observed serial intervals of length t by $s(t)$. Let $\lambda(u)$ denote the number of secondary transmissions occurring u days after the onset of the primary case (including the days for $u < 0$). We then obtain the expected number of serial intervals of length Δt by the convolution equation:

$$E[s(\Delta t)] = \sum_i \sum_{k(i)} \int_{-x}^{\Delta t} \lambda(u)f(t_{k(i)} - (t_i + u))\mathrm{d}u, \quad (2)$$

where $k(i)$ denotes the secondary cases infected by case i. The basic idea of eqn (2) can now be linked to the method of back-calculation in HIV/AIDS [24].

Since we assume the incubation period distribution to be known, estimates of disease-age-specific infectiousness can be obtained in a non-parametric fashion [25] assuming a step function model for $\lambda(u)$:

$$\left. \begin{array}{ll} \lambda(u) = \rho_1 & \text{for} \quad -5 \leqslant u < 0 \\ \lambda(u) = \rho_2 & \text{for} \quad 0 \leqslant u < 3 \\ \lambda(u) = \rho_3 & \text{for} \quad 3 \leqslant u < 6 \\ \lambda(u) = \rho_4 & \text{for} \quad 6 \leqslant u < 9 \\ \lambda(u) = \rho_5 & \text{for} \quad 9 \leqslant u < 38 \\ \lambda(u) = 0 & \text{otherwise.} \end{array} \right\} \quad (3)$$

Further details are given in the Appendix.

Figure 2a shows the estimated daily frequency of secondary transmissions with corresponding 95% confidence intervals (CI) based on the profile likelihood. Between 3 and 6 days after onset of fever, which roughly corresponds to the time immediately after the onset of rash, the daily frequency of secondary transmissions was highest, yielding 20·6% (95% CI 15·1–26·4) of the total number of secondary transmissions. Expected cumulative frequencies of secondary transmissions before onset of fever and up to 3 days after onset of fever were 2·6% and 23·7%, respectively. The obtained estimates suggest that 91·1% of all secondary transmissions occurred up to 9 days after onset of fever. Figure 2b compares the observed and predicted serial intervals. Figure 2b confirms a relatively good overall agreement between predicted and observed data, but the χ^2 goodness-of-fit test revealed a significant deviation between the observed and expected serial intervals (see legend for Fig. 2), which is largely attributable to outliers in the observed data.

The significance of estimating disease-age-specific infectiousness of smallpox lies in the practical implications both on the population and the individual level. Previous studies frequently assumed a deterministic (fixed) incubation period of 11 or 12 days to infer the date of smallpox infection [26, 27]. As isolation of cases is critically important to control smallpox [9, 10], it is crucial to evaluate the time at which isolation can be delayed. Our estimates are also useful for the evaluation of surveillance containment measures (i.e. ring vaccination) where the disease-age-specific infectiousness plays a key role [28]. Our study implies that isolation could be extremely effective if performed before onset of rash and that delayed isolation of symptomatic cases could still be effective if performed within a few days after onset of rash. The
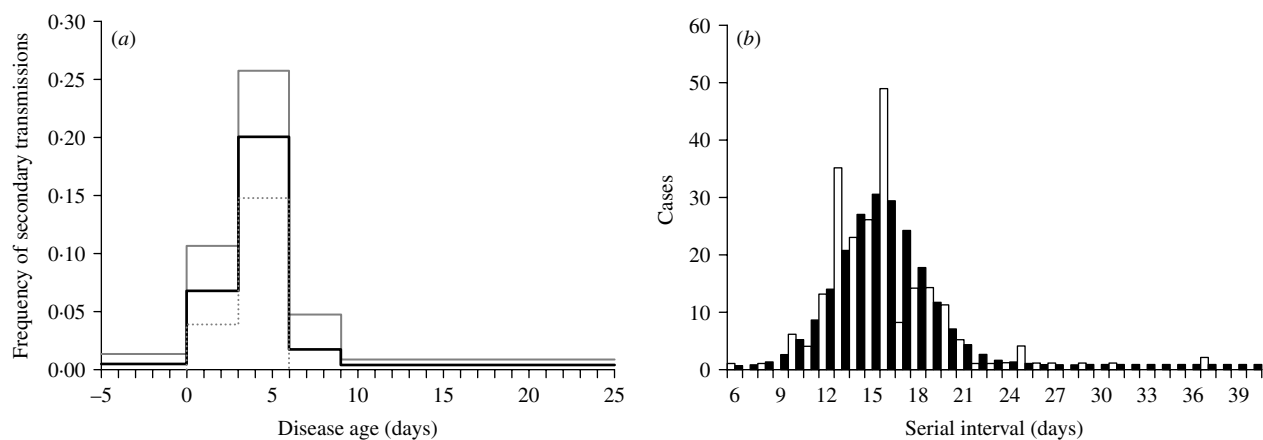
Fig. 2. Frequency of secondary transmissions and predicted serial intervals. (*a*) Expected daily frequency of secondary transmissions with corresponding 95% confidence intervals (CI) (—, upper 95% CI; —, expected; ....., lower 95% CI). Disease age $t=0$ denotes the onset of fever. (*b*) Observed (□) and predicted (■) daily counts of the serial intervals ($n=223$). The χ^2 test revealed significant deviation between the observed and predicted values ($\chi^2_5 = 20 \cdot 9$, $P < 0 \cdot 01$). When assessing the goodness-of-fit of our models, using the χ^2 statistic, we divided the serial interval into ten groups ($\leqslant 11$, 12–13, 14–15, 16–17, 18–19, 20–21, 22–23, 24–25, 26–27 and $\geqslant 28$ days).

highest infectivity immediately after onset of rash is consistent with previous suggestions based on observational experience [12, 13, 29].

This paper proposes a simple method to estimate the disease-age-specific infectiousness. So far, similar information for an acute infectious disease has only been obtained in a rigorous study based on longitudinal data of household transmissions of influenza using Markov chain Monte Carlo [30]. Our study shows that the incubation period and precise data on who infected whom enables the estimation of the infectiousness without specific settings such as households. The proposed method is extremely useful as far as independence of transmission during the course of disease can be assumed for diseases with acute course of illness. (Note: this assumption is not recommended for slowly progressing diseases such as HIV/AIDS [18].) However, the probability of transmission tends to be biased by various factors: our small sample of serial intervals may have been influenced by local factors such as differences in contact behaviour and mobility of cases. This partly explains why several outliers are observed in the serial intervals (Fig. 2*b*). For example, Fig. 3 shows the total number of cases who contributed to the generation of secondary cases (based on serial intervals) and the subset of those who were admitted to the hospital before onset of the secondary cases in Brighton during 1950–1951 [20]. Other data sources hardly provided us with similar information. The impact of isolation, which could not be explicitly incorporated in this
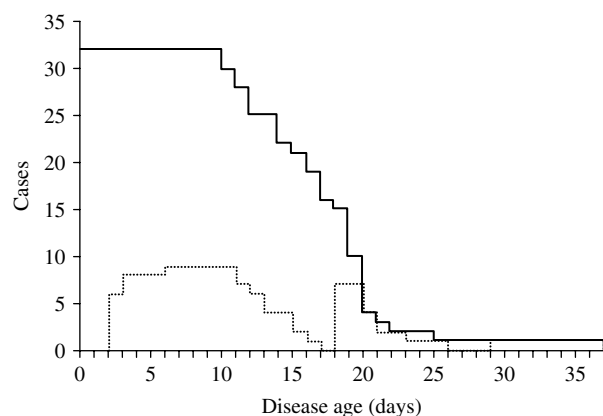
Fig. 3. Number of infectious cases and those isolated among them by disease age. The solid line shows how many index cases were followed by secondary cases t days after onset of symptoms and had to be considered as still being infectious in Brighton (1950–1951). The dotted line shows how many of these index cases were hospitalized on the given disease age.

study due to small sample sizes, can be assessed by looking at the area under the dotted line (Fig. 3). Thus, the infectiousness several days after appearance of rash may have been underestimated. This could partly explain a disagreement with a previous study [31] in which the infectiousness during the prodromal period was estimated as 8·2% of the overall transmission potential. In conclusion, this study showed a simple back-calculation of infectiousness by disease age, based on the known distribution of the

incubation period distribution and the transmission network.

APPENDIX: Likelihood-based approach

Since the serial interval is given with a daily precision, we consider the convolution eqn (2) in discrete time:

$$E(s_t) = \sum_i \sum_{k(i)} \sum_{u=-x}^{t} \lambda_u f_{t_{k(i)} - (t_i + u)}, \qquad (A\ 1)$$

where f_k is the probability that the incubation period has length k days. Assuming that λ_u is generated by a non-homogeneous Poisson process, resulting in s_t serial intervals of length t, the likelihood function, which is needed to estimate λ_u, is proportional to

$$\prod_{t=-5}^{\infty} (E(s_t))^{r_t} \exp(-E(s_t)), \qquad (A\ 2)$$

where r_t denotes the daily counts of the serial interval (Fig. 2b). The maximum-likelihood estimates of λ_u were obtained by minimizing the negative logarithm of eqn (A 2). The 95% CI was determined using the profile likelihood.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **Greenwood M.** On the statistical measure of infectiousness. *Journal of Hygiene (Cambridge)* 1931; **31**: 336–351.
2. **Hope-Simpson RE.** Infectiousness of communicable diseases in the household (measles, chickenpox, and mumps). *Lancet* 1952; **2**: 549–554.
3. **Dietz K.** The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research* 1993; **2**: 23–41.
4. **Diekmann O, Heesterbeek JAP.** *Mathematical Epidemiology of Infectious Diseases: model building, analysis and interpretation.* New York: Wiley Series in Mathematical and Computational Biology, 2000.
5. **Anderson RM, May RM.** *Infectious Diseases of Humans: dynamics and control.* Oxford: Oxford University Press, 1991.
6. **Becker NG.** *Analysis of Infectious Disease Data.* London: Chapman & Hall, 1989.
7. **Shiboski SC, Jewell NP.** Statistical analysis of the time dependence of HIV infectivity based on partner study data. *Journal of the American Statistical Association* 1991; **87**: 360–372.
8. **Koopman JS, et al.** The role of early HIV infection in the spread of HIV through populations. *Journal of Acquired Immune Deficiency Syndrome and Human Retrovirology* 1997; **14**: 249–258.
9. **Fraser C, et al.** Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences USA* 2004; **101**: 6146–6151.
10. **Eichner M.** Case isolation and contact tracing can prevent the spread of smallpox. *American Journal of Epidemiology* 2003; **158**: 118–128.
11. **Chu CM, et al.** Viral load distribution in SARS outbreak. *Emerging Infectious Diseases* 2005; **11**: 1882–1886.
12. **Downie AW, et al.** Studies on the virus content of mouth washings in the acute phase of smallpox. *Bulletin of the World Health Organisation* 1961; **25**: 49–53.
13. **Sarkar JE, et al.** Virus excretion in smallpox. 1. Excretion in the throat, urine, and conjunctiva of patients. *Bulletin of the World Health Organisation* 1973; **48**: 517–522.
14. **Haydon DT, et al.** The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 2003; **270**: 121–127.
15. **Lloyd-Smith JO, et al.** Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; **438**: 355–359.
16. **Nishiura H, et al.** Transmission potential of primary pneumonic plague: time-inhomogeneous evaluation based on historical documents of the transmission network. *Journal of Epidemiology and Community Health* 2006; **60**: 640–645.
17. **Fine PE.** The interval between successive cases of an infectious disease. *American Journal of Epidemiology* 2003; **158**: 1039–1047.
18. **Wallinga J, Teunis P.** Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology* 2004; **160**: 509–516.
19. **Laidlaw SIA, Horne WA.** Smallpox outbreak in Glasgow, 1950. *Medical Officer* 1950; **83**: 187–190.
20. **Cramb R.** Smallpox outbreak in Brighton, 1950–51. *Public Health* 1951; **64**: 123–128.
21. **Lyons J, Dixon CW.** Smallpox in the Industrial Pennines, 1953. *Medical Officer* 1953; **90**: 293–300, 307–310.

22. **De Sario V.** Field investigation of an outbreak of smallpox at Bawku, Ghana: May–October, 1967. *East Africa Medical Journal* 1970; **47**: 524–535.

23. **Mukherjee MK, Sarkar JK, Mitra AC.** Pattern of intrafamilial transmission of smallpox in Calcutta, India. *Bulletin of the World Health Organisation* 1974; **51**: 219–225.

24. **Brookmeyer R, Gail MH.** A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *Journal of the American Statistical Association* 1988; **83**: 301–308.

25. **Becker NG, Watson LF, Carlin JB.** A method of non-parametric back-projection and its application to AIDS data. *Statistics in Medicine* 1991; **10**: 1527–1542.

26. **Rao AR.** *Smallpox.* Bombay: The Kothari Book Depot, 1972.

27. **Mack TM.** Smallpox in Europe, 1950–1971. *Journal of Infectious Diseases* 1972; **125**: 161–169.

28. **Kretzschmar M, et al.** Ring vaccination and smallpox control. *Emerging Infectious Diseases* 2004; **10**: 832–841.

29. **Fenner F, et al.** *Smallpox and its Eradication.* Geneva: World Health Organization, 1988 (http://whqlibdoc. who.int/smallpox/).

30. **Cauchemez S, et al.** A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data *Statistics in Medicine* 2004; **23**: 3469–3487.

31. **Eichner M, Dietz K.** Transmission potential of smallpox: estimates based on detailed data from an outbreak. *American Journal of Epidemiology* 2003; **158**: 110–117.